M.SC DATA ANALYTICS & TECHNOLOGIES

DATA ANALYSIS ON A SUPER STORE DATASET USING EXCEL

MODULE: DATA SCIENCE

CODE: DAT 7006

STUDENT ID: 2312897

MODULE TUTOR: DR ANCHAL GARG

TABLE OF CONTENTS

INTRODUCTION

The Dataset used in this project was that of a superstore containing orders made by different customers from the year 2014 to 2017. Analysis was done on the dataset to gain insights and make effective decision-making. The tool used to perform this task was Microsoft Excel. With Excel, I was able to perform steps that included the following

I. Data Cleaning

II. Data Processing

III. Data Analysis

IV. Data Visualization

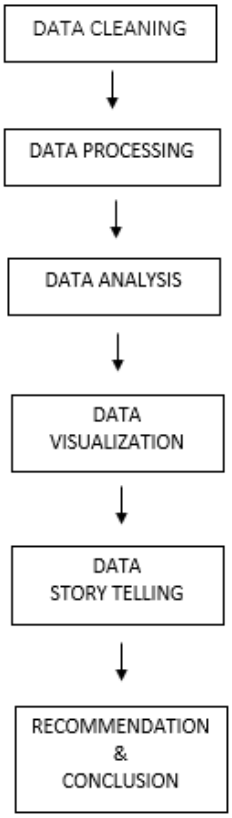

Figure 1-1: Flowchart

1.1 DATA CLEANING

Data cleaning is the process of identifying and correcting errors, inconsistencies, and inaccuracies in a dataset to improve its quality and reliability for analysis or modeling. This typically involves tasks such as removing duplicate entries, handling missing values, correcting data formatting issues, and detecting and rectifying outliers.

In the dataset, it was discovered that 1 duplicate data was found. All columns were converted to the correct data format and all missing values were solved by dropping the rows that contained the missing values.

In the column named **Category,** it was discovered that some rows had the category spelled wrongly as it was seen that Furniture was spelled furtiure and Technology spelled as Techonology. This was solved by using the **Find all and replace all** feature present in Excel.

1.2 DATA PROCESSING

Data processing refers to the manipulation and transformation of raw data into a more useful format or structure. It involves tasks such as sorting, summarizing, aggregating, filtering, and transforming data to extract meaningful insights or support decision-making.

In the dataset, the following columns were created by applying the correct formulas to be able to extract more meaningful insights

I. Sales/Quantity

II. Cost price/Quantity

III. Profit/Quantity

1.3 DATA ANALYSIS AND VISUALIZATION

Data analysis involves interpreting data to extract insights, identify patterns, and make informed decisions. Data visualization is the graphical representation of data to visually communicate insights, trends, and patterns. It involves creating charts, graphs, and other visualizations to help understand complex data sets and present findings clearly and understandably. Data visualization enhances understanding and aids in decision-making by making data more accessible and intuitive to interpret.

Research questions were asked and answers were discovered

1.3.1 RESEARCH QUESTION 1

Why is the country experiencing some losses? The relationship was drawn between the discounts given on products and the profit gotten from each order. On analysis, it was discovered that products with higher discounts tend to run at losses for the company.

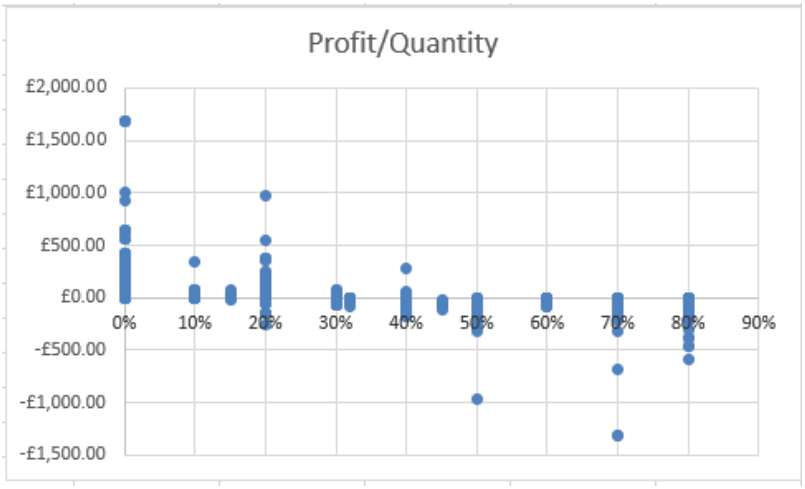

Figure 1-2: Profit/Discount Relationship

1.3.2 RESEARCH QUESTION 2

Which category has the most sales? On analysis, it was discovered that Technology had the most sales, with total sales of £836,330.83, while Furniture had £741,591.58. This enables the superstore which products are most patronized enabling them to make decisions.

Sum of Sales	Category			
	Furniture	Office Supplies	Technology	Grand Total
Total	£741,591.58	£718,932.13	£836,330.83	£2,296,854.55

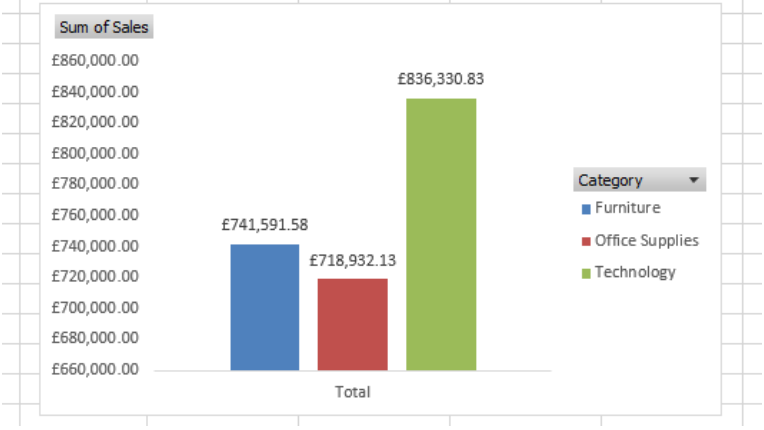

Figure 1-3: Sales across different categories

1.3.3 RESEARCH QUESTION 3

Categories of the returned orders? The returned orders were present in another table, with the use of INDEX and MATCH formulas, I was able to match the order ID with that of the main orders table, and with the use of the index, I got the category it belonged to. This helped the superstore know which category experienced more returned orders to know the next step to take. From the analysis, it was discovered that Office supplies had the most returned orders.

Count of Order ID (Return)	Returned Category			
	Furniture	Office Supplies	Technology	Grand Total
Total	73	175	48	296

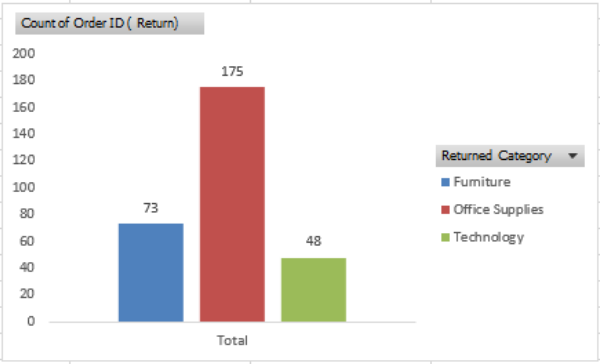

Figure 1-4: Returned orders

1.3.4 RESEARCH QUESTION 4

Profits made across different categories likewise across different regions? This played a vital role in providing information on which category and region amounted to the most profit for the superstore this will assist in improving the marketing strategy across the different regions.

Region	(All)			
Sum of Profit	Category			
	Furniture	Office Supplies	Technology	Grand Total
Total	£ 18,206.11	£ 122,456.33	£ 145,477.93	£ 286,140.37

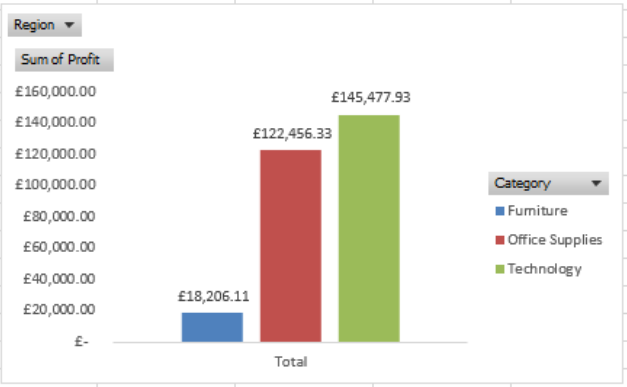

Figure 1-5: Profits made across regions

1.3.5 RESEARCH QUESTION 5

To find out the shipping mode across different states? Performing this task enables the superstore know which shipping mode is most employed in order to make necessary planning in order to get orders delivered swiftly.

State	(All)				
Count of Row ID	Ship Mode				
	First Class	Same Day	Second Class	Standard Class	Grand Total
Total	1538	543	1947	5967	9995

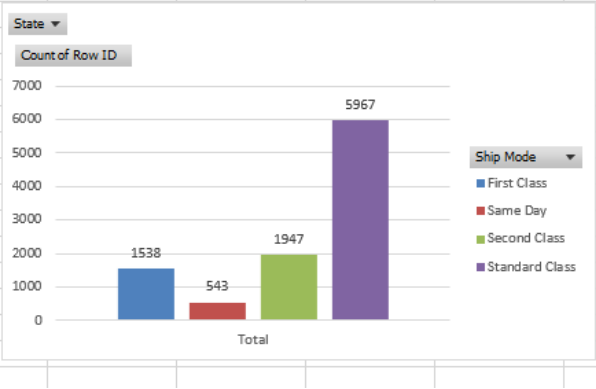

Figure 1-6: Shipping mode employed across states

1.3.6 ASKED QUESTION 1

Here, the work towards finding out the discount given to the customer named Alejandro Grove. I used a filter to filter the customers' names and provide the information of Alejandro Grove. On application, it was discovered that he had benefited from only a 20% discount on an order he had made on the 27th of March, 2017 for furniture.

Figure 1-7: Discount for Alejandro Grove

1.3.7 ASKED QUESTION 2

To find out the total sales on each date? With the aid of a pivot table, analysis was done to determine the total sales made on each day with the use of filter for each day. For example, the image below shows the total sales made for the 7th of January, 2014.

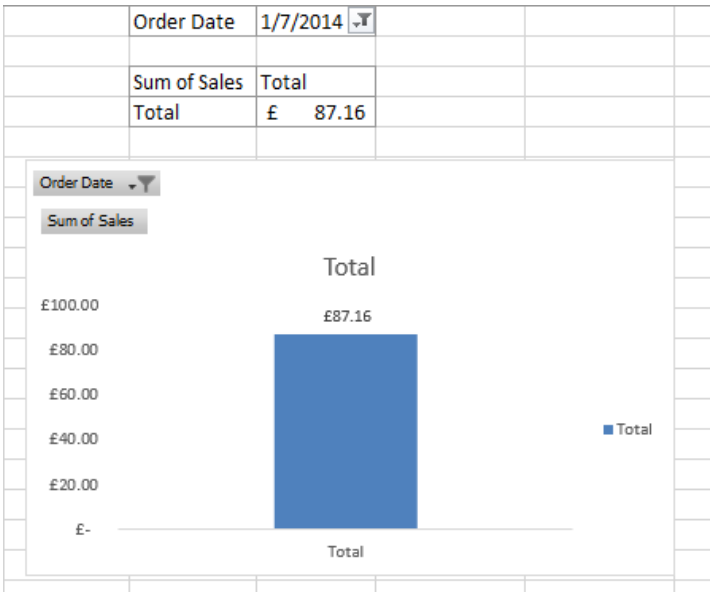

Figure 1-8: Total sales made daily

1.3.8 ASKED QUESTION 3

To find the total sales year-wise? Same approach is followed as discussed in Asked Question 2, the only difference is we apply the filter to filter by year and not days. For example, the image below shows the total sales made for the year 2014.

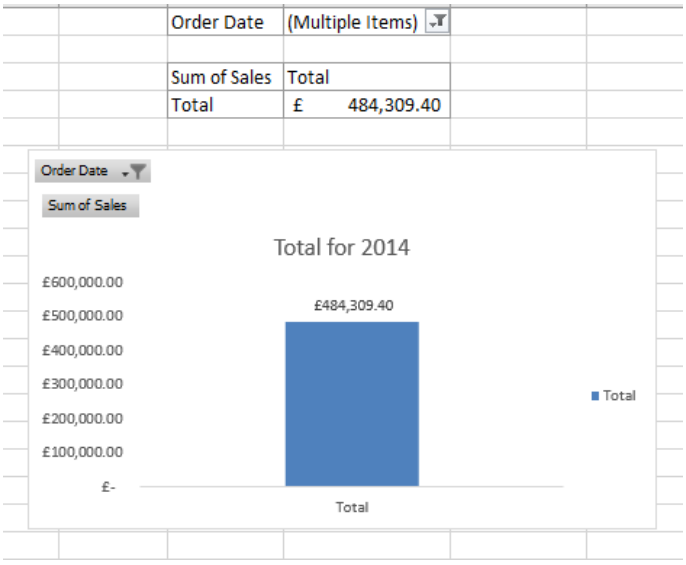

Figure 1-9: Total Sales made yearly

1.4 RECOMMENDATION AND CONCLUSION

From analysis it can be seen that generally the superstore is performing well economically but will require some adjustments in order to increase profits and sales likewise of which include

I. Reduction of high discounts placed on products,

II. Evaluating on the required discount to be placed on some products in order to realize profits,

III. Focus on preparing questionnaires for customers that returned order to know the reason behind such action and how to take necessary steps to prevent reoccurrence.